PROMA: AN ONTOLOGY BASED APPLICATION FOR EDUCATION/INDUSTRY COLLABORATION

Houssem Gasmi\textsuperscript{1,2}, Abdelaziz Bouras\textsuperscript{1}

\textsuperscript{1} College of Engineering, Qatar University (QATAR)
\textsuperscript{2} DISP Lab, Lyon 2 University (FRANCE)

Abstract

As the main supplier of workforce to the industry, higher education is increasingly challenged for not being abreast with the digital revolution and somehow disconnected from the industry. Competency-based education was therefore developed to address this issue and bridge the gap between what the university is producing and the requirements of the industry. To analyze this gap, tools need to be developed that assists in the analysis process. This paper focuses on proposing a system that models the competencies of occupations in the industry and higher education curricula and assist in matching profiles from the two domains. The different concepts in the domain are modelled as a semantic web ontology and the profile matching is performed by an inference engine. In addition to the profile matching, the system calculates a matching score using the Analytic Hierarchy Process (AHP) method.

Keywords: Ontology, Competency Modeling, Matching System.

1 INTRODUCTION

The ever increasing advances in the industry necessitates a close collaboration between the industry and the universities. Higher education programs need to be kept up to date to cope with the industry's high qualified labor needs to ensures students are better prepared for the needs of industry and increase their marketability. Traditionally, the main goal of the academic institutions is the production of knowledge and transferring it to its students. However, this role is not adequate anymore and universities need to prepare the student for a successful career when he joins his workplace. Increasingly over the years, several universities and academic institutions in different countries have incorporated competency frameworks in their curricula design activities. Competency frameworks are one of the educational tools used to ‘vocationalise’ academic courses. Their main objective is to describe in detail the skills, knowledge, attitudes, and typical activities of a job type that is prepared by a training course. During the seventies, competency-based education started to appear in the United states and later in other western countries. This approach was the result of the growing criticisms towards traditional education which was considered as more and more disconnected from the social evolution of that time, especially changes within the workplaces [1]. It has been used initially to reform upper-secondary vocational curricula and more recently academic institutions started to employ it as a tool to reform their courses and keep them up to date [2], [3]. Several local initiatives have been proposed to formalize the definition of competencies such as O*NET in the United States and ROME in France.

The aim of this paper is to present a profile matching application that models the domain knowledge as a semantic web ontology [4]. Ontologies has been used to solve issues of interoperability between different domains [5]. Different parties usually have different concepts and needs that is the result of their different perspective. Ontologies make reuse much easier by avoiding the wasted efforts of term translation, where the shared understanding of a term in a certain domain does not lead to mismatched concepts and different definitions. The objective of the proposed application is to establish an ontological relationship between the competency requirements of Qatar market and the higher education learners profile to ensure a continuous alignment between student profiles and the industry. The educational and industry profiles are represented as O*NET competency framework profiles. The system consists of two major parts, an ontology that models the domain concepts and a service that calculates the degree of matching between industry profiles and educational profiles using the AHP method [6]. This paper focuses on describing the application architecture, the functionality it provides, an overview of the ontology and score calculation.

This paper is organized as follows. In section two we review related work. Section three provides an overview of the application architecture. Section four introduces the different stakeholders and their
interactions with the application. Section five provides an overview of the ontology design and the logic behind profile matching. Section six shows how the system calculates the matching score between a university curriculum and an occupation using the AHP method. Section 7 concludes the paper.

2 RELATED WORK

In this section, we review the related work in profile matching and ontology mapping by selecting the papers that are more relevant to our domain of research. Many recruitment tools for measuring the suitability of candidates for a job have appeared in recent years. However, there has been little research investigating the semantic matching systems between the concrete needs of the industry and educational outcomes of universities. Thus, most of the developed tools suffer from inadequate matching of candidates with job requirements. To the best of our knowledge, no previous published work has applied a profile matching system integrating semantic information related to the industry occupations and competencies generated by education curricula in order to evaluate the relevance between the profiles.

Several attempts were made to automate the recruitment process [7] using recommendation systems that match the job with the suitable candidate. The goal of these recommendation systems is to speed-up and increase the efficiency of the recruitment process. Several recommendation techniques have been applied such as content based filtering, rule-based filtering, collaboration filtering and hybrid filtering [8]. These recommendation systems combine the data collected from the user with related data collected from other sources and generate a recommended list of items for the user. These systems find the match between people skills and job offer descriptions taking into consideration the preferences of the recruiters and the interests of the candidates. Some recommendation systems such as CASPER [9] are based on user interactivity and employ collaborative filtering techniques to make recommendation for users based on the preferences of other users that have similar profiles.

The PROSPECT system is proposed in [10] to select candidates for recruitment. It extracts selection criteria from resumes such as education, competencies and relevant experience in each. [11] suggests a standardized format for resumes to improve the effectiveness of candidates filtering and selection. A mobile based recommendation system was proposed by [12] catered for career services in universities. It helps at matching recruiting companies with graduated students at low cost and focuses on two-sided profile matching based on preference lists for further recommendations. Several approaches are used by the recruitment systems like relevance feedback, natural language processing [13], semantic matching, machine learning [14], and the automatic representation of resumes based on a standard format using the analytic hierarchy process [15].

Domain ontologies have become a mainstream knowledge representation tool in many recent applications. Previous related decision-support works in the educational domain have used context ontology-based approach [8], [16]. Recommendation systems have also applied ontologies as a key part in their efficient filtering techniques. In order to improve the vocational based education, we want to provide in this paper a semantic matching application based on ontological models that allows us to better capture, analyze and use relevant semantic information for the study and analysis of the gap between education and the industry. Consequently, the proposed application helps universities to satisfy both companies and students. We describe the proposed application in the next section.

3 PROPOSED APPLICATION

The proposed application is a desktop application and has two main responsibilities: the first responsibility is matching between an education profile and an industry profile and state whether the two profiles match. The education profile can be either a student transcript or a university curriculum. The structure of the two profiles is similar, the only difference is that the student profile contains his/her grades which are used as an additional criterion in the matching score calculation. The industry profile represents an occupation we want to match against a student transcript or check whether a university curriculum fulfills the occupation requirements. The profiles are represented as instances of classes in the PROMA ontology which will be covered in more details in a later section. The second responsibility of the application is calculating the matching score for matched profiles using the AHP method taking into account the matching criteria defined in the system. (Figure 1. PROMA Architecture) depicts the profile matching application architecture. The application is comprised of four main components: the PROMA ontology, the inference engine, the Profile Matching Service, and the Matching Score Calculation Service.
The PROMA ontology defines the concepts in the domain and the relationship between them in addition to the matching rules. The ontology is designed to require defining the minimal information possible and infer the rest using an inference engine which is a piece of software able to infer logical consequences from a set of asserted facts [17]. The ontology is persisted in the file system although more sophisticated ontology containers can be used as the application needs scaling.

The Profile Matching Service (PMS) component is the heart of the application, it loads the unclassified profiles from the ontology and invokes an external logic inference engine called Pellet. The language used to read profiles from the ontology is SQWRL which is a query language designed specifically to read information from ontologies [18]. An example to load student transcripts from the ontology is shown below:

```
Student(?s) ^ hasTranscript(?t) ^ consistsOfEnrolment(?t, ?e)^
  is-An-EnrolmentForCompetency(?e, ?c) ^ hasGrade(?e, ?g)
→ sqwrl:select(?s, ?t, ?c, ?g)
```

This query will return the list of students along with their full transcript which contains the gained competencies and the grades they obtained for each competency. SQWRL is accessed through the SQWRL API (Application Programming Interface) which is a Java-based API and is the interface between the PMS component and the ontology. Once the ontology is loaded, the inference engine executes a set of SWRL rules [19] which define the matching rules to populate the ontology with more inferred knowledge and perform the profile matching which results eventually in the classification of all the education profiles to either Matched or Not Matched with their associated industry profiles. The PMS component then updates the profiles in the ontology by marking all the processed ones to prevent them from being loaded in subsequent processing. The matched profiles are then converted into an object model and passed to the Matching Score Calculation Service component. This component reads all the matched profiles and calculates the matching score for each one using the AHP method. The profiles are then updated with the calculated score and the ontology is updated.
The application has two user interfaces, a conventional graphic user interface and a set of RESTful APIs [20] hosted in a web server to provide the direct communication with the ontology and issue queries about the different aspects of the domain like university curricula and occupations available. The graphical user interface (Figure 2. PROMA User Interface) displays the list of occupations defined in the ontology and the list of university curricula. It allows the user to drill down to list the competencies required by an occupation and also to list the courses of a university curriculum and their learning outcomes. Each learning outcome has a competency type (skill, ability, and knowledge) associated with it. Once an occupation and a curriculum are selected, the user can click on the calculate matching score button which starts the calculation and displays the result.

![Figure 2. PROMA User Interface](image)

4 SCENARIOS

As a collaboration tool, the main application usage is as an analysis tool by collaboration committees from the education and industry sectors to analyze the discrepancies between a curriculum and an occupation. The system can equally be used as a recruitment tool by industry companies to select candidates for recruitment. As depicted in (Figure 3. PROMA Use case diagram), the main actors of the system are as follows:

- **Curriculum Designer:** this person defines the courses which constitute a university curriculum and for each course defines its learning outcomes. He is also responsible for evaluating the curriculum against related occupations in the industry based on the matching results of the application. The Curriculum Designer is actually part of an internal committee that is responsible for designing and updating the university curricula.

- **Student:** a student can evaluate his profile against an occupation. Based on the program he selects and the grades he obtained, he gets a matching score describing how fit he is for the selected occupation.

- **Recruiter:** this person defines an occupation in the system. The definition must include in detail the list of competencies required and their requirements. He can calculate the matching score for student applicants against occupation openings.

- **Domain Expert:** this person is responsible for mapping between a university curriculum and an occupation. He must have industrial experience and a strong academic background. The Domain expert is part of a committee that consists of people from academia and people from the industry that is responsible for the evaluation of university curricula.
The main use cases of the application are the following:

- **Define Curriculum**: defines the courses which constitute a university curriculum and their learning outcomes.
- **Evaluate Curriculum**: calculates a matching score for the curriculum against occupations in the industry.
- **Define Occupation**: defines an occupation and its required competencies.
- **Evaluate Student Profile**: calculates a matching score for a student profile against an occupation.
- **Define Mapping**: maps learning outcomes of a curriculum to the competencies required by an occupation. This mapping is the basis of the profile matching. (Education/Industry Profiles Mapping)

5 CONCEPTUAL MODEL

The core component of the system is the profile matching ontology which represents the different aspects of the domain model and performs the matching between profiles using logical rules executed by an inference engine. The model consists mainly of classes and objects properties. Classes represent groups of individuals that share the same properties and are arranged in a hierarchy to model sub type relationships. The hierarchy of the domain classes for the PROMA model is shown in (Figure 5. PROMA Classes Hierarchy). Properties are ontology concepts that represent the relationship between class individuals. For example, the Student class is related to other classes such as Application, Transcript using the properties hasApplication and hasTranscript respectively as shown (Figure 4. PROMA Class definition). The tool used to build the model is called Protégé. It is one of the most popular free tools for ontology modeling and it comes with a built in inference engine. The ontology is populated only with the necessary instances and derives as much information as possible using the inference engine.
The profile matching model can be logically divided into the following sub-models:

a) Common model
   - Competency: represents the central class in the model and is shared among all the sub models. It has three sub classes: ability, skill, and knowledge. A competency can be an outcome of a course and gained by a student taking a course, or is a requirement of an occupation.

b) Education model: the education profile consists of the following classes:
   - Study Plan: represents a curriculum and consists of a list of courses taught by a department in the university.
   - Course: a course of study that is normally recognized for credit towards the granting of an approved degree. Each course should result in a list of learning outcomes gained by the student.
   - LearningOutcome: The learning outcomes of a course. A learning outcome is equivalent to one or more competencies.

c) Student model: the student profile consists of the classes defining the Education profile, in addition to the following classes:
   - Student: represents a student in a university.
Transcript: represents the list of courses taken by the student and the grades he obtained.

Enrolment: it is part of the transcript and is a combination of a course taken by a student and the grade he obtained in that course.

Grade: the grades for the courses are represented with the letters A, B, C, D, and F.

d) Application model: the application sub-ontology consists of the following classes:

- Application: represents a student application for an occupation.
- MatchingOutcome: represents the outcome of the student application for an occupation.

e) Occupation model:

- Occupation: represents an occupation in the industry.
- Competency Requirement: each competency required by an occupation has a name, an importance level, and a required competency level.
- Importance Level: represents how important the competency is for the recruiter and it can take the following values: Required, Preferred, Desired.
- Competency Level: specifies the required level for a competency in an occupation and can take the following values: Knowing, Capable, Competent.
- MappingLevel (Relevance): captures the degree of relevance between a curriculum competency and an occupation competency and take the following values: Weak, Related, Strong.

The profile matching logic in the ontology is expressed as a set of rules expressed in the Semantic Web Rule Language (SWRL) (Figure 6. PROMA Ontology Rules). SWRL is used as a tool to infer more knowledge from the ontology. For example, the relation between an Occupation and a Competency (represented by the object property requiresCompetency) is not directly asserted in the ontology definition but inferred using rules (see Figure 6. PROMA Ontology Rules. These rules are also the basis for the matching logic which is performed as follows:

- Given an occupation profile and an education profile, the list of competencies of both profiles are matched against each other using the following methods:
  - Competency matching by name: this method looks for competencies in both profiles that have the same name and match them (see Figure 6. PROMA Ontology Rules.
  - Competency matching by defined mapping: it is usually the case that two competencies represent the same competency but have different names. In this case, a mapping is defined in the model and is used in the matching process to match these types of competencies (see Figure 6. PROMA Ontology Rules.

- The two profiles are matched if all the competencies required by the occupation profile are matched by competencies in the education profile either by name or by explicit mapping definition.

---

**Rules:**

```prolog
\[ A \]
\[ \text{Occupation(?o), hasCompetencyRequirement(?o, ?r), forCompetency(?r, ?c) } \rightarrow \text{ requiresCompetency(?o, ?c)} \]
\[ \text{NotMatched(?no) } \rightarrow \text{ hasMatchingResult(?no, Not_Full)} \]
\[ \text{Student(?s), hasTranscript(?s, ?t), consistsOfEnrolment(?e, ?r), inCourse(?e, ?c), generatesLearningOutcome(?c, ?o), equivalentToCompetency(?o, ?c) } \rightarrow \text{ isAnEnrolmentForCompetency(?enrolment, ?competency)} \]
\[ \text{hasMatchingOutcome(?o, matchingResult), Student(?s), SameAs (?o1, o2), forOccupation(?a, ?o), hasGainedCompetency(?s, ?c1), hasApplication(?s, ?a), requiresCompetency(?a, ?c2) } \rightarrow \text{ hasMatchedCompetency(?matchingResult, ?c1)} \]
\[ \text{Application(?a), forOccupation(?a, ?o1), hasMatchingOutcome(?a, ?o2), ((is-A-OccupationForApplication some Application) and (requiresCompetency only (belongsToMatchingOutcome some MatchingOutcome)))(?o2), SameAs (?o1, ?o2) } \rightarrow \text{ hasMatchingResult(?o2, Full)} \]
```

**Figure 6. PROMA Ontology Rules**

6925
6 MATCHING SCORE CALCULATION USING AHP

Once profiles are matched, the next step is to calculate the level of matching between them. Profile matching score calculation involves calculating the matching score for each matched competency in the profile, then the total matching score for the profile is found by taking the average of the competency matching scores. The AHP method is used as an analysis tool to evaluate the level of matching between competencies against a set of three criteria. The criteria of matching are: Importance Level, Mapping Level, and the Competency Level which applies to student profiles. The values each of these criteria can take has been described in the previous section.

As per the AHP method, the starting point is to select the analysis goal. In our case, the analysis goal is the calculation of a matching score between two profiles. The second step is to calculate the weight of each criterion based on its importance for a domain expert or a recruiter. The combination of the resulting criterion weights is called the weight vector. The criteria weight vector is then used to calculate the matching score between competencies. Suppose we have the profiles shown in (Education/Industry Profiles Mapping) below. The industry profile consists of the competencies under the Industrial Competency column and the education profile consists of the competencies under the Education Competency column. The two profiles final matching score is the average of the matching score of each row in the table.

<table>
<thead>
<tr>
<th>Industrial Competency</th>
<th>Relevance</th>
<th>Importance</th>
<th>Education Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming in Java</td>
<td>Related</td>
<td>Required</td>
<td>Programming 101</td>
</tr>
<tr>
<td>Network Security</td>
<td>Strong</td>
<td>Desired</td>
<td>Computer Networks</td>
</tr>
<tr>
<td>Linux Administration</td>
<td>Weak</td>
<td>Preferred</td>
<td>Operating Systems</td>
</tr>
</tbody>
</table>

7 CONCLUSION

Studying the gap between higher education outcomes and the industry needs is an important endeavor that is worth investigating. However, developing efficient and accurate solutions for measuring that gap and exploring it proved to be a challenging task to all stakeholders both in education and industry. The application presented in this paper showed how the ontologies and the semantic web technologies can be used to adequately capture the domain concepts to perform the required gap analysis. The ontology is considered as a mean that facilitates the communication between the different stakeholders and serves as a foundation for future collaboration that leads to reaching a solid competency model. As a future work, we would like to make the application more automated in terms of profile definition and the mapping of profiles which is an activity performed by a domain expert. These activities are labor intensive and time consuming.

ACKNOWLEDGEMENT

This publication was made possible by NPRP grant # NPRP 7-1883-5-289 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES


